УДК 004.932

О.І. Гетьманець, О.О. Гагарін
Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Україна
пр. Перемоги, 37, м. Київ, 03056

АЛГОРИТМ ПЕРЕТВОРЕННЯ ВІДЕОПОТОКУ ДЛЯ ПРИСКОРЕННЯ ТА ОПТИМІЗАЦІЇ ПОШУКУ ПЕРЕМІЩЕННЯ ОБ'ЄКТУ

O.I. Getmanets, O.O. Gagarin
National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Ukraine
37, Peremohy av., Kyiv, 03056

ALGORITHM OF VIDEO STREAM TRANSFORMATION FOR ACCELERATION AND OPTIMIZATION OF SEARCH OF OBJECT TRANSMISSION

У статті запропоновано алгоритм перетворення кадрів безперервного потоку відео для прискорення та оптимізації пошуку переміщення об'єкта в режимі реального часу. Сучасні широко розповсюджені системи відеонагляду виконують задачу пошуку руху у відеопотоці, але вони не завжди стійкі до неякісних вхідних даних і сповільнюються у разі обробки великих об'ємів даних за короткі проміжки часу. Запропонований у статті алгоритм розбиває процес пошуку переміщення об'єкта на етапи, на кожному з яких виконується перетворення кадрів відеопотоку, що покращує процес пошуку першого зрушення об'єкта та подальшого відслідковування його переміщення.

Ключові слова: відеопотік, колірна модель, сегментація, монохромне зображення, бінарізація.

In the article the algorithm of continuous video frame transformation stream for acceleration and optimization of search of object moving in the real-time mode is offered. Video surveillance systems fulfill the task of searching for movements in a video stream, but they are usually not resistant to inferior input data and are slowing down in the case of processing large volumes of data at short intervals. The algorithm that is proposed divides this process to stages, each of which performs the transformation of the video stream frames, which improves the process of finding the first shift of the object and further tracking its movement.

Keywords: video stream, color model, segmentation, monochrome image, binarization.

Вступ

Аналіз зображень відеопотоку в найзагальнішому вигляді являє собою порівняння ряду послідовних (у часі) цифрових зображень сцени спостереження з метою реєстрації різного роду змін (найчастіше – переміщення об'єктів). Вирішення такого завдання широко розповсюджене в системах безпеки та відеонагляду.

Постановка проблеми

Процес пошуку переміщення об'єкта у відеопотоці має три основні проблеми:

1. Дані надаються та мають бути аналізовані в режимі реального часу.
2. Алгоритми аналізу мають бути стійкими до неякісних вхідних даних (погане освітлення, перенасичений кольорами фон тощо).
3. Необхідно аналізувати великі об'єми даних.

Огляд існуючих рішень

Окрім існуючих рішень у вигляді різних систем відеонагляду, новітні дослідження представлені Інститутом інженерів з електротехніки та електроніки [1, 6].

Зокрема, у 2014 році на Міжнародній конференції з робототехніки та біоміметики (IEEE International Conference on Robotics and Biomimetics) група учасників презентувала власний алгоритм перетворень зображень для знаходження

положення об'єкта в кадрі [1]. У даній роботі було запропоновано новий підхід до точної сегментації об'єкта на основі комбінації регіональної сегментації та оціночної орієнтації, яка обчислюється глобальним альбедо-контрастом проектованих і захоплених зображень. У процесі дані розподілені на фоновий та об'єктний блоки (регіони), які окремо обробллюються і результати цих перетворень об'єднуються в загальний результат виділення шуканого об'єкта.

У результаті запропонований авторами алгоритм працював швидко і був стійким до вхідних даних різної якості, але не був пристосований для роботи в реальному часі, так як кожна ітерація пошуку вимагала повного циклу перетворень обох вхідних зображень.

Запропонований алгоритм

Існуючі рішення слабо адаптовані для вирішення трьох визначених проблем одночасно, тому в даній статті пропонується покращений алгоритм перетворення відеопотоку. Далі наведено покроковий опис алгоритму, який покращує показники процесу пошуку рухомих об'єктів у відеопоптоці.

У реальному часі задача пошуку переміщення об'єкта має дві послідовні складові: захоплення руху та відслідковування рухомого об'єкта. Перша частина вимагає точності, а друга – швидкості роботи. Тому доцільно розділити алгоритм перетворень на два окремих кроки, що вирішують проблему роботи в режимі реального часу.

Перший етап (визначення першого моменту зміщення об'єкта) виконується через певні фіксовані проміжки часу. При цьому вимоги до розміру проміжків досить лояльні, адже в реальних умовах немає критичної різниці у виявленні зрушення об'єкта на 5-ій чи 15-ій секунді. Головне – уникнути помилкового виявлення. Отже, на цьому етапі має вирішуватись проблема стійкості до неякісних вхідних даних.

З математичної точки зору будь-яке зображення можна представити у вигляді трьох незалежних величин ⟨x, y, I⟩, де пара цілочисельних значень ⟨x, y⟩ описує геометричне положення даного пікселя в площині зображення, а величина I характеризує його яскравість або інтенсивність в даній точці площини. Для випадку кольорових зображень величина I має три складові. Наприклад, для найрозповсюдженішої колірної моделі RGB (Red, Green, Blue) це буде значення червоної, зеленої та синьої складової кожного пікселя. При цьому, складність аналізу кожного зображення прямо пропорційна його розміру, а час залежить також від кількості кольорів для порівняння. Так, якщо в напівтоновому (в сірих відтінках) зображенні піксель може мати 256 відтінків яскравості, то в разі кольорового зображення піксель може мати один з 256 відтінків яскравості кожного з трьох кольорів, що робить кількість можливих кольорів рівним 2563 = 16 777 216. Логічно було б припустити, що для аналізу більш доцільно відкинути кольори або працювати лише з конкретною колірною складовою, але експерименти доводять, що чорно-білі кадри не обов'язково будуть давати найкращий результат.

У наступній таблиці (Таблиця 1) показано відсотки помилкових виявлень руху та рухомих об'єктів для різних за розмірами та якістю вхідних даних [2]. З таблиці видно, що в середньому найкращі результати показує аналіз мультфільмів, головна відома особливість яких – невелика кількість кольорів (особливо їх відтінків у межах одного об'єкта) та чіткі межі об'єктів. Обробка чорно-білого відео також дає непогані результати, але вони поступаються мальованим зображенням.

Таблиця 1. Помилкові виявлення

Метод	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*
Мультфільми	10%	10%	31%	18%	0%
Кольорове відео	33%	11%	33%	33%	27%
Чорно-біле відео	23%	-	25%	-	12%

Методи оцінювання стійкості до неякісних вхідних даних:

1. Палітра розбивається по інтенсивності.
2. Палітра розбивається на RGB-паралелепіпеди.
3. Квадродерева і палітра розбивається по інтенсивності.
4. Квадродерева і палітра розбивається на RGB-паралелепіпеди.
5. Палітра розбивається по інтенсивності, застосовується ускладнена формула відстані і вирівнювання гістограм.

Отже, перед тим як переходити до порівняння двох кадрів, необхідно поділити зображення на логічні ділянки та зменшити таким чином вагу впливу неякісних даних у кожній ділянці. Для цього використовують кольорову сегментацію, тобто виокремлення частини зображення з конкретним кольором. Оскільки реальні об’єкти найчастіше представлені у вигляді не конкретного кольору, а групи близько розташованих кольорів (відтінків), то модель RGB з чітким розподіленням кольорів для цієї задачі не підходить. Алгоритми сегментації найбільш пристосовані для роботи з HSV (Hue, Saturation, Value – колірний тон, насиченість, міра яскравості) – модель, що забезпечує можливість явного задання необхідного відтінку. Перехід з простору RGB у простір HSV дозволяє уникнути помилки в колірній сегментації, тобто «втрати» меж об’єкта у випадку неякісних вхідних даних, наприклад, коли зображення «засвічене» або навпаки затемнене.

У процесі досліджень була проаналізована велика кількість алгоритмів сегментації, при цьому основним критерієм була можливість їх застосування в режимі реального часу для безперервного потоку кадрів. Далі представлені короткий опис найбільш оптимальних методів та результати їх оцінювання для вирішення заявленої задачі.

Текстурний метод базується на аналізі чотирьох параметрів: середнього значення, дисперсії, орієнтації та масштабу. Відносна фіксованість цих значень у часі спрощує ітераційний аналіз зображень, але якщо текстура змінює свої характеристики залежно від масштабу, текстурні ознаки, інваріантні щодо масштабу, можуть не існувати взагалі. Тоді використання текстур для опису об’єктів на різних відстанях стає складним і затратним завданням.

На відміну від текстурного аналізу, метод Отса використовує лише гістограму розподілу значень яскравості пікселів растрового зображення. За умови використання HSV моделі, цей метод дає найшвидший результат для сильного спрощення кольорів при сегментації.

Метод k-середніх є методом кластерного аналізу, який у процесі прагне мінімізувати сумарне квадратичне відхилення точок кластерів від центрів цих кластерів. Головний недолік методу – велика його залежність від заздалегідь визначених кластерів (їх розмірів, розміщення та кількості). У випадку зображень, які перенасичені об’єктами, цей метод досить часто надає помилковий результат.

З основною проблемою методу k-середніх успішно справляється вдосконалений метод нечіткої c-середньої кластеризації. Замість однозначної відповіді на питання, до якого кластера відноситься об'єкт, він визначає ймовірність того, що об'єкт належить до того чи іншого кластера. Це покращує результат обробки, але, з іншого боку, потребує більших затрат пам'яті, а іноді також додаткової подальшої обробки.

У Таблиці 2 наведено середні значення часу виконання (у мілісекундах) кожного описаного методу (для методів кластеризації – кластери задаються розбиттям зображення на рівну кількість частин) для різних за якістю та розмірами вхідних кадрів [3].

Таблиця 2. Час виконання сегментації зображень

Метод	*Великі вимоги до точності (велика кількість відтінків у результаті)*	*Низькі вимоги до точності (незначна кількість відтінків у результаті)*	*Перенасичене об'єктами та кольорами зображення*
Текстурний аналіз	0.55031 мс	0.10686 мс	4.16039 мс
Отса	0.58775 мс	0.0564 мс	3.87603 мс
k-середніх	0.57884 мс	0.14781 мс	3.62995 мс
c-середніх (з урахуванням пост-обробки)	0.62491 мс	0.06806 мс	4.12004 мс

З даних таблиці видно, що, в середньому, найкращий результат отримано в разі використання методу Отса та k-середніх. Враховуючи, що перетворення має бути стабільним до неякісних вхідних даних, пропонується використання методу Отса як методу з найкращим співвідношенням швидкість-якість. Таким чином вирішується проблема стійкості до неякісних вхідних даних.

Після сегментації вже виділені локальні ділянки за загальною колірною ознакою, а отже колір більше значення не має і можливий перехід до палітри відтінків сірого, тобто до монохромного зображення. При цьому кольорова модель значення не має, а найоптимальніший метод перетворення представлений у формулі (1), де i, j – це координати пікселя, Gray – отримане значення відтінка сірого, R, G та B – червона, зелена та синя складові кольору кожного пікселя відповідно.

$$Gray_{ij} = 0.3 * R_{ij} + 0.59 * G_{ij} + 0.11 * B_{ij} \quad (1)$$

Фіксування першої розбіжності (першого моменту руху) зазвичай відбувається на даному кроці, коли зображення спрощені, але все ще надають повну інформацію про об'єкти. Для оптимізації процесу найчастіше використовують блокові алгоритми пошуку, тобто попереднє розбиття зображень на частини для подальшого пошуку руху в кожному окремому блоці. Таке рішення дозволить реалізувати обробку паралельно, що добре підходить для великої кількості даних, які необхідно аналізувати в режимі реального часу.

У роботі [4] було проаналізовано різні підходи до пошуку рухомих об'єктів. Основним критерієм досліджень була точність результату, чого вимагає перший крок запропонованого алгоритму.

Статистичний підхід має переваги в простоті реалізації та в низьких обчислювальних витратах, але методи його реалізації мають низьку якість у разі зміни геометричних форм об'єктів і мають слабку стійкість до поганих вхідних даних, що може бути критичним дефектом саме у застосуванні до кадрів відеопотоку.

Параметричний підхід більш гнучкий для реальних умов зйомок, в тому числі до шумів на зображеннях та у випадку перенасиченого об'єктами та кольорами фону, але він потребує значних обчислювальних ресурсів.

Так як на даному етапі вимоги до точності переважають над іншими вимогами, запропоновано використання одного з найбільш стійких параметричних методів – методу відслідковування за точковими особливостями (Feature Tracking). Він стійкий до поганих вхідних даних і має низьку ймовірність помилкового детектування. Алгоритм, що полягає у пошуку зміщення Δx точки x у межах ω (де I – безпосередньо зображення, а t – складова часу), представлено у формулах (2) – (5).

$$E = \sum_{\omega(x)}(I(x+\Delta x,t) - I(x,t+\Delta t))^2 \quad (2)$$

$$I(x+\Delta x) \approx I(x) + \Delta x\frac{\partial}{\partial x}I(x) \quad (3)$$

$$0 = \frac{\partial}{\partial \Delta x}E \approx \sum_{\omega(x)}\frac{\partial}{\partial \Delta x}\left(I(x,t) + \Delta x\frac{\partial}{\partial x}I(x,t) - I(x,t+\Delta t)\right)^2 \quad (4)$$

$$\Delta x = \left[\sum_{\omega(x)}\left(\frac{\partial I}{\partial \Delta x}\right)^T((I(x,t) - I(x,t+\Delta t))\right]\left[\sum_{\omega(x)}(\frac{\partial I}{\partial \Delta x})^T(\frac{\partial I}{\partial \Delta x})\right]^{-1} \quad (5)$$

Таким чином завершується перший етап – етап пошуку зрушення об'єкта. На другому етапі відбуватиметься відслідковування руху вже знайденого об'єкта, отже змінюються вимоги до алгоритмів – тепер необхідне швидке опрацювання великих об'ємів даних для вирішення останньої вказаної проблеми.

Після першого знаходження рухомого об'єкта немає сенсу знову перебирати все зображення, так як зазвичай більша його частина не має інформаційної цінності для аналізу. Тоді для пришвидшення обчислень можливо перетворити монохромне зображення на однобітове бінарне. Це – ще більш спрощена модель, у якій значення «кольору» може бути лише 1 або 0, при цьому 0 прийнято вважати фоном (у даному випадку – нерухомою частиною зображення, яка ігнорується), а 1 – це власне об'єкт для відслідковування.

Процес бінарізації базується на принципі порогових значень, тобто розбиття зображення на дві області, одна з яких містить всі пікселі зі значенням нижчим деякого порогу, а інша містить усі пікселі зі значенням вищим цього порогу. Враховуючи те, що рухомий об'єкт вже знайдено, бінарізацію запропоновано здійснювати не за принципом розділення по конкретному кольору, а з виділенням об'єкта та фону, які можуть складатись з кількох (навіть частково схожих) кольорів (формула (6)). Пороги бінарізації будуть визначатися на основі дельти зображень (де дельта – рухомий об'єкт або його частина). Тобто, пропонується проводити бінарізацію в декілька етапів, для кожного кольору, по якому було визначено приналежність частини зображення до рухомого об'єкта.

$$f'(m,n) = \begin{cases} 0, f(m,n) \geq t_1; \\ 1, t_1 < f(m,n) \leq t_2; \\ 0, f(m,n) > t_2; \end{cases} \quad (6),$$

де f(m,n) – кожен піксель зображення, а t_1, t_2 – крайові значення для конкретного кольору (відтінку сірого) визначеного рухомого об’єкта.

Так як після вищеописаних перетворень вже є чітке розділення об’єкта та фону, для подальшого відслідковування руху можна використати статистичний підхід, зокрема метод просторової фільтрації [5]. Принцип просторових алгоритмів полягає в застосуванні спеціальних операторів до кожної точки вихідного зображення. Операторами виступають прямокутні або квадратні матриці, так звані маски. Найчастіше маска являє собою невеликий двовимірний масив, а методи оптимізації, що базуються на такому підході, часто називають обробкою по масці або фільтрацією по масці. У застосуванні цього методу для пошуку руху маскою вважають перший кадр на вході, який порівнюється з наступним кадром. Результат порівняння знаходять за формулою (7).

$$R = w(-1,-1)f(x-1,y-1) + w(-1,0)f(x-1,y) + \cdots + w(0,0)f(x,y) + \cdots + w(1,0)f(x+1,y) + w(1,1)f(x+1,y+1) \quad (7),$$

де w(x,y) – маска, f(x,y) – зображення, x та y – координати точок.

На рисунку 1 зображено вхідне зображення для другого етапу (монохромне та розбите на блоки) та результат подальшого відслідковування руху на наступній серії кадрів (вже бінарних). Білим кольором відмічено рухомі ділянки, а чорно-білі клітинки – нерухомий фон.

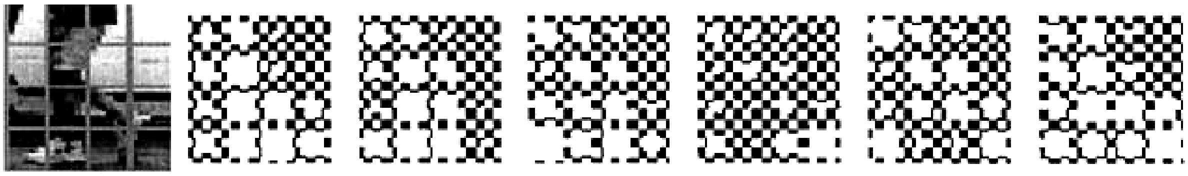


Рис. 1. Результат пошуку руху на бінарізованих зображеннях

Таким чином швидко відбуваються ітерації другого етапу, коли відслідковується вже знайдений рухомий об’єкт і на вхід подається велика кількість зображень через малі проміжки часу.

Висновки

Запропонований алгоритм ділиться на два логічних етапи. На першому етапі виконуються сегментація та монохромізація, що дає змогу точно знайти зрушений з місця об’єкт в режимі реального часу при неякісних вхідних даних. На другому етапі відбувається бінарізація зображення, в результаті якої рухомий об’єкт логічно відділяється від фону і відслідковування його руху відбувається швидко навіть для великого об’єма вхідної інформації.

Запропонований алгоритм вирішує зазначені проблеми для задачі пошуку рухомого об’єкта у відеопотоці. Його включено до розробок охоронних систем, які проектуються.

Література

1. Мінг Х. Альбедо-контрастна сегментація рук у системі проектор-камера в режимі реального часу / Х. Мінг, С. Джун, В. Вей. // Міжнародна конференція з робототехніки та біоміметики. – 2014.
2. Деякі підходи до організації змістовного пошуку зображень та відеоінформації / Н.С. Байгарова, Ю.А. Бухштаб, Н.Н. ЄвтеЄва, Д.А. Корягін. // Препрінт ІПМ ім. М.В.Келдиша РАН. – 2002. – №78.
3. Комбінований підхід для посилення та сегментації маммограми з використанням модифікованого методу нечіткого с-методу у вейвлет-області / С. Срівастава, Н. Шарма, С. Сін, Р. Срівастава. // Журнал медичної фізики. – 2014. – №39.
4. Методи пошуку руху у відеопослідовностях / М.Н. Фаворська, А.И. Пахірка, А.С. Шилов, М.В. Дамов. // Вісник Сибірського державного університету науки і технологій імені академіка М.Ф. Решетньова. – 2009.
5. Гонсалес Р.С. Цифрова обробка зображень / Р.С. Гонсалес, Р.Є. Вудс., 2001. – (2).
6. Ні Ю. Адаптивна модель для пошуку швидкої оцінки відповідності блоків / Ю. Ні, К. Ма. // Транзитна обробка зображень. – 2002. – №11. – С. 1442–1448.

Literatura

1. Minh Kh. Al'bedo-kontrastna sehmentatsiya ruk u systemi proektor-kamera v rezhymi real'noho chasu / Kh. Minh, S. Dzhun, V. Vey. // Mizhnarodna konferentsiya z robototekhniky ta biomimetyky. – 2014.
2. Deyaki pidkhody do orhanizatsiyi zmistovnoho poshuku zobrazhen' ta videoinformatsiyi / N.S. Bayharova, Yu.A. Bukhshtab, N.N. YevteYeva, D.A. Koryahin. // Preprint IPM im. M.V.Keldysha RAN. – 2002. – #78.
3. Kombinovanyy pidkhid dlya posylennya ta sehmentatsiyi mammohramy z vykorystannyam modyfikovanoho metodu nechitkoho c-metodu u veyvlet-oblasti / S. Srivastava, N. Sharma, S. Sin, R. Srivastava. // Zhurnal medychnoyi fizyky. – 2014. – #39.
4. Metody poshuku rukhu u videoposlidovnostyakh / M.N. Favors'ka, A.Y. Pakhirka, A.S. Shylov, M.V. Damov. // Visnyk Sybirs'koho derzhavnoho universytetu nauky i tekhnolohiy imeni akademika M.F. Reshetn'ova. – 2009.
5. Honsales R.S. Tsyfrova obrobka zobrazhen' / R.S. Honsales, R.Ye. Vuds., 2001. – (2).
6. Ni Yu. Adaptyvna model' dlya poshuku shvydkoyi otsinky vidpovidnosti blokiv / Yu. Ni, K. Ma. // Tranzytna obrobka zobrazhen'. – 2002. – #11. – S. 1442–1448.

RESUME

O.I. Getmanets, O.O. Gagarin

Algorithm of video stream transformation for acceleration and optimization of search of object transmission

The proposed algorithm fulfills the task of finding a moving object in a video stream in a step-by-step solution for solving three main problems: real-time operation, resistance to poor-quality input data and processing large volumes of data.

Analyzed existing solutions for finding the difference between video frames work fast and are resistant to incoming data of different quality, but are not adapted for real-time operation, since each iteration of the search requires a complete cycle of transformations of both input images.

In the first stage, segmentation and monochromization are performed, which allows precisely to locate the object shifted from the site in real time with low-quality input data. Among the existing segmentation algorithms, the best result is obtained by using the Otsu's and k-medium methods. The transformation must be stable to poor-quality inputs, so it is proposed to use the Otsu's method as the method with the best rate-quality ratio. This way the problem of resistance to low-quality input data is solved. After segmentation and monochromization the search for the first shift of the object is proposed. Since at this stage the accuracy requirements prevail over other requirements, one of the most stable parametric methods – the Feature Tracking method is suggested. It is resistant to poor input data and has a low probability of false detection.

At the second stage, the binaryization of the image occurs. Since after the above-described transformations there is already a clear separation of the object and the background, a statistical approach can be used to further track the movement, in particular the spatial filtration method. As a result, the moving object logically separates from the background and tracking its movement occurs quickly even for a large amount of input information.

The proposed algorithm solves these problems for the task of finding a moving object in a video stream and is included in the development of security systems that will be projected in the future.

Надійшла до редакції 31.10.2017

